



**Improved production of biopharmaceuticals by site-specific
cleavage of fusion proteins expressed in *Escherichia coli*.**

**A thesis submitted for the degree of Doctor of Philosophy at the
University of Adelaide, Department of Molecular Bioscience**

**Adam Charlton,
B.Biotech. (Hons)
Flinders University, South Australia**

May 2007

TABLE OF CONTENTS

ABSTRACT	x
DECLARATION	xii
ACKNOWLEDGEMENTS	xiii
ABBREVIATIONS.....	xiv

CHAPTER 1: LITERATURE REVIEW

1.1	FUSION PROTEINS	1
1.2	FUSION PARTNER ROLES	7
1.2.1	<i>The fusion protein strategy to prevent intracellular proteolysis.....</i>	
1.2.2	<i>The fusion protein strategy for the expression of insoluble protein.</i>	9
1.2.3	<i>The fusion protein strategy for the soluble expression of protein ...</i>	12
1.2.4	<i>The fusion protein strategy for the provision of an affinity tag</i>	15
1.2.5	<i>Conclusion.....</i>	16
1.3	THE FUSION PROTEIN STRATEGY IN THE PRODUCTION OF INSULIN-LIKE GROWTH FACTORS.....	17
1.	<i>The Met-pGH fusion partner for IGF production</i>	21
1.4	CLEAVAGE OF FUSION PROTEINS.....	23
1.4.1	<i>Chemical cleavage of fusion proteins.....</i>	24
1.4.2	<i>Proteolytic cleavage of fusion proteins.....</i>	25

1.4.3	<i>Site-specific proteases of the mammalian blood clotting cascade.</i>	31
1.4.4	<i>Mammalian gastric site-specific serine proteases</i>	33
1.4.5	<i>Viral site-specific cysteine proteases</i>	36
1.4.6	<i>Engineered site-specificity: H64A subtilisin</i>	38
1.4.7	<i>Strategies to correct suboptimal cleavage</i>	40
1.4.8	<i>Conclusion</i>	43
1.5	α -LYTIC PROTEASE	43
1.5.1	<i>Rational α-Lytic protease mutants</i>	45
1.5.2	<i>Combinatorial α-Lytic protease mutant libraries</i>	47
1.5.3	<i>The other subsites of α-Lytic protease</i>	48
1.5.4	<i>Conclusion</i>	50
1.6	CLEAVAGE OF IGF FUSIONS	50
1.6.1	<i>Cleavage of IGF-I fusions</i>	52
1.7	CLEAVAGE OF THE MET-PGH(1-11)VN-IGF-I FUSION PROTEIN WITH α -LYTIC PROTEASE: THIS PROJECT	55

CHAPTER 2: CULTURE AND QUANTITATION OF THE α -LYTIC PROTEASE MUTANT LIBRARIES

2.1	INTRODUCTION	58
2.2	MATERIALS	66
2.2.1	<i>General reagents & materials</i>	66
2.2.2	<i>Purified α-Lytic protease standard</i>	66

2.2.	<i>Bacterial strains</i>	67
2.2.4	<i>Bacterial culture media</i>	\
2.2.5	<i>General solutions</i>	68
2.3	METHODS.....	68
2.3.1	<i>Expression of α-Lytic protease mutants</i>	68
2.	<i>Colourimetric peptide assay for PragA9 activity</i>	69
2.3.3	<i>Protease culture on skim milk agar</i>	70
2.	<i>Fluorescence polarisation conjugate preparation</i>	70
2.3.5	<i>Fluorescence Polarisation assay</i>	71
2.	<i>SDS-PAGE of FTC-casein conjugate</i>	71
2.3.7	<i>Enzyme Immunoassay</i>	72
2.4	RESULTS.....	73
2.4.	<i>Culture of mutants</i>	73
2.4.2	<i>Synthesis of conjugate</i>	74
2.4.3	<i>FP standard curve</i>	74
2.4.4	<i>FP conjugate SDS-PAGE</i>	75
2.4.5	<i>Quantitation of protease activity in culture supernatants by FP</i> ..	76
2.4.6	<i>EIA standard curve</i>	79
2.4.7	<i>Quantitation of protease activity in culture supernatants by EIA</i> ..	79
2.5	DISCUSSION.....	80
2.6	CONCLUSION.....	87

CHAPTER 3: SCREENING THE α -LYTIC PROTEASE MUTANTS FOR DESIRABLE

SPECIFICITY

3.1	INTRODUCTION.....	88
3.2	MATERIALS	91
3.2.1	<i>General reagents & materials.....</i>	91
3.2.2	<i>FP peptide sets</i>	92
3.2.3	<i>Discrete peptides.....</i>	92
3.2.4	<i>General solutions.</i>	93
3.2.5	<i>α-Lytic protease purification buffers.....</i>	93
3.3	METHODS	93
3.3.1	<i>FP peptide screening assay</i>	
3.3.2	<i>Purification of shortlisted mutants.....</i>	94
3.3.3	<i>Cleavage of discrete peptides</i>	95
3.3.4	<i>Mass Spectrometry of peptide cleavages.....</i>	96
3.4	RESULTS	96
3.4.1	<i>FP screening peptide design,</i>	96
3.4.2	<i>Peptide screening assay</i>	98
3.4.3	<i>α-Lytic protease shortlist selection and purification</i>	102
3.4.4	<i>Cleavage of long discrete peptides.</i>	103
3.4.5	<i>Mass Spectrometry of peptide cleavage.....</i>	110
3.5	DISCUSSION.....	111
3.6	CONCLUSION	120

CHAPTER 4: CLEAVAGE OF MET-PGH(1-11)VN-PAPM-IGF-I

4.1	INTRODUCTION.....	121
4.2	MATERIALS.....	125
4.2.1	General reagents & materials.....	125
4.2.2	General solutions.....	126
4.2.3	Purified PragA9 α -Lytic protease.....	126
4.2.4	Fusion protein substrates.....	127
4.3	METHODS.....	127
4.3.1	Conduct of cleavage reactions.....	127
4.3.2	Analysis of cleavage reactions by HPLC.....	127
4.3.3	SDS-PAGE of cleavage reactions.....	128
4.3.4	Mass Spectrometry of peptide cleavages.....	129
4.3.5	S-Carboxymethylation reactions.....	129
4.4	RESULTS.....	129
4.4.1	HPLC analysis of variations in cleavage conditions.....	129
4.4.2	SDS-PAGE of Met-pGH(1-11)VN-PAPM-IGF-I cleavage.....	148
4.4.3	Mass spectrometry of Met-pGH(1-11)VN-PAPM-IGF-I cleavage	149
4.4.4	Cleavage of S-Carboxymethylated Met-pGH(1-11)VN-PAPM-IGF-I.	151
4.4.5	HPLC analysis of Met-pGH(1-11)VN -IGF-I cleavage reactions.....	154
4.5	DISCUSSION.....	156

4.6	CONCLUSION	163
-----	------------------	-----

CHAPTER 5: MANIPULATION OF LEADER PEPTIDE AND CLEAVAGE SITE STRUCTURE

5.1	INTRODUCTION.....	
5.2	MATERIALS	169
5.2.1	<i>General reagents and materials</i>	169
5.2.2	<i>Molecular biology reagents</i>	170
5.2.3	<i>Oligonucleotides and Plasmids</i>	170
5.2.4	<i>Bacterial strains</i>	171
5.2.5	<i>Bacterial culture media</i>	171
5.2.6	<i>Protein production and analysis</i>	172
5.2.7	<i>Protein synthesis assay</i>	174
5.2.8	<i>General solutions</i>	174
5.3	METHODS	175
5.3.1	<i>Generation of the Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I expression construct</i>	175
5.3.2	<i>Generation of the Met-pGH(1-11)VN-WGSGPAPM-IGF-I expression construct</i>	176
5.3.3	<i>Agarose gel electrophoresis</i>	178
5.3.4	<i>Plasmid preparation and DNA sequencing</i>	178
5.3.5	<i>Trilal induction</i>	179
5.3.6	<i>Fusion protein production</i>	179
5.3.7	<i>Harvesting and washing of Inclusion bodies</i>	

5.3.8	<i>Inclusion body dissolution and oxidative refolding of novel Met-pGH(1-11)VN-PAPM-IGF-I constructs</i>	180
5.3.9	<i>Fusion protein purification</i>	181
5.3.10	<i>Fusion protein cleavage</i>	182
5.3.11	<i>Preparative cleavage product separation</i>	182
5.3.12	<i>L6 Myoblast protein synthesis assay</i>	183
5.3.13	<i>SDS-PAGE and Western blot analysis</i>	184
5.3.14	<i>Analysis of cleavage reactions by HPLC</i>	185
5.3.15	<i>Mass Spectrometry of proteins</i>	185
5.4	RESULTS	185
5.4.1	<i>Generation of Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I expression plasmid</i>	185
5.4.2	<i>Generation of Met-pGH(1-11)VN-WGSGPAPM-IGF-I expression plasmid</i>	187
5.4.3	<i>Expression of Met-pGH(1-11)VN-WGSGPAPM-IGF-I and Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I</i>	189
5.4.4	<i>Downstream processing of Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I</i>	195
5.4.5	<i>Downstream processing of Met-pGH(1-11)VN-WGSGPAPM-IGF-I</i>	197
5.4.6	<i>Cleavage of Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I</i>	199
5.4.7	<i>Cleavage of Met-pGH(1-11)VN-WGSGPAPM-IGF-I</i>	203

5.4.8	<i>Separation of Met-pGH(1-11)VN-WGSGPAPM-IGF-I cleavage products</i>	
5.4.9	<i>Mass spectrometry of Met-pGH(1-11)VN-WGSGPAPM-IGF-I cleavage products</i>	207
5.4.10	<i>Optimisation of Met-pGH(1-11)VN-WGSGPAPM-IGF-I cleavage</i> ..	213
5.4.11	<i>L6 myoblast protein synthesis</i>	221
5.5	DISCUSSION	226
5.5.1	<i>Expression</i>	226
5.5.2	<i>Folding</i>	226
5.5.3	<i>Selection rationale</i>	228
5.5.4	<i>Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I isomer 1 cleavage</i> ...	229
5.5.5	<i>Met-pGH(1-11)VN-WGSGPAPM-IGF-I isomer 2 cleavage</i>	231
5.5.6	<i>Reaction pathway</i>	232
5.5.7	<i>des(68-70) truncations</i>	233
5.5.8	<i>Protein synthesis</i>	235
5.5.9	<i>Met-pGH(1-11)VN-[F¹⁶E/M¹³E]-PAPM-IGF-I</i>	238
5.5.10	<i>Process yields</i>	239
5.6	CONCLUSION	240

CHAPTER 6: GENERAL DISCUSSION

6.1	INTRODUCTION.....	241
6.2	SUGGESTED IMPROVEMENTS TO THE STRATEGY.....	245
6.3	OTHER OBSERVATIONS	247
6.3.1	<i>Cleavage at P₁' Proline.....</i>	247
6.3.2	<i>The influence of structure on cleavage.....</i>	248
6.4	FUTURE DIRECTIONS.....	250
6.4.1	<i>Short term goals.....</i>	250
6.4.2	<i>Longer term goals</i>	252
	REFERENCES	254

ABSTRACT

The recombinant expression of heterologous proteins in microorganisms, such *Escherichia coli*, is often improved by producing the protein of interest translationally linked to another, often unrelated, protein giving rise to a "fusion protein" construct. For many applications it is desirable or imperative to separate the extraneous material from the protein of interest. An increasingly popular approach to this task is the use of site-specific endoproteases to excise the protein product. A number of commercially available site-specific proteases exist, but many are not capable of generating an authentic N-terminus for the product, display unsatisfactory specificity leading to adventitious cleavage of the product, or they are unsuitable for an industrial process.

Mutants of the serine protease α -Lytic protease have been shown to satisfy many of the criteria for an industrially suitable protease and have been applied to the cleavage of some important fusion proteins used in the production of members of the Insulin-like Growth Factor (IGF) family. Lacking from these examples, however, is any viable proteolytic solution for the liberation of human IGF-I from fusion proteins. This has been primarily attributed to the Proline bearing N-terminal tripeptide sequence of this protein, which is known to be refractory to the activity of many site-specific proteases.

It has been suggested that, in the generation of two combinatorial mutant libraries of α -Lytic protease, the preference for amino acids C-terminal to the cleavage site may have been altered. It is the purpose of this work to first determine if such an alteration has been made in any of the mutants so as to

allow cleavage immediately before the N-terminus of human IGF-I, and then to task the lead mutant(s) to the cleavage of the full-length fusion protein. All members of the two mutant libraries were cultured and their activity confirmed and quantified against a generic β -casein substrate in a high-throughput assay. A second high-throughput technique was then employed to query the mutant proteases for their ability to catalyse proteolysis at the required sequence in a peptide model. Finding that many mutants appeared successful at this task, the findings were verified on a longer peptide model of the cleavage site.

Initially the yields achieved by cleavage of the full-length IGF-I fusion protein by a lead candidate mutant α -Lytic protease were not sufficient to satisfy the requirements of an industrial process, despite alteration of the reaction conditions. However, the insight gained from these reactions could be applied to the redesign of the protein structure around the intended site of cleavage, significantly improving site-specific proteolysis. The IGF-I generated by this cleavage has been shown to be bioequivalent to commercial reference standard to cultured mammalian cells and the yield of this process is approximately 5-fold improved over the existing cleavage system.